Text S1. Supplemental Materials and Methods, and supplemental information about mating type analyses.

Contents of Text S1

1. Supplemental Materials and Methods

a. Growth conditions, preparation of nucleic acids and RNA-seq libraries

For genomic DNA isolation, a culture of 30 mL YPD medium was cultivated for 16 h at 28°C and 180 rpm. After centrifugation, (5 min at 3,000 g) cells were resuspended in distilled water (0.5 mL), centrifuged and resuspended in lysis buffer (2% Triton X-100, 1% sodium dodecyl sulfonate (SDS), NaCl (100 mM), Tris/HCl (10 mM, pH 8), EDTA (1 mM)). 300 mg glass beads, phenol (100 µL), chloroform (96 µL), and isoamyl alcohol (4 µL) were added, and the mixture was rigorously shaken for 3 min. After centrifugation (10 min, 10,000 g, 4°C), the upper layer was transferred to another tube containing the same mixture of phenol, chloroform and isoamyl alcohol and the extraction was repeated for three times. A solution of RnaseA (5 µL, 10 mg/mL) was added and the solution incubated for 5 min at 37°C. The extraction with phenol, chloroform and isoamyl alcohol was repeated. Then, an additional extraction was done with chloroform (96 µL) and isoamyl alcohol (4 µL) and the upper layer

was added to a solution of ammonium acetated (4 M, 0.1 volumes) and ethanol (96%, 2 volumes) and centrifuged. After washing with ethanol (70%), the pellet was dried and resuspended in distilled water (50 µL).

For transcriptome analysis, *T. oleaginosus* was grown on five different media, media compositions are given in Text S1-Table S1. 50 mL medium were inoculated in a 250 mL flask and incubated at 28°C and 180 rpm. The samples for transcriptome analysis for YPD were taken after one day, for NAG and XYL after two days, for NLM and PLM after five days. 10 mL culture were centrifuged (5 min, 3,000 g) and the cells were resuspended in AE buffer (400 µL, sodium acetate (50 mM), EDTA (10 mM), pH 5.3). SDS (40 µL) and phenol (550 µL) were added and the mixture was shaken vigorously. After incubation at 65°C for 4 min, the mixture was cooled on ice for 15 min. After centrifugation for 2 min (15.000 g, 4°C) the aqueous phase was transferred to a pre-cooled aqueous solution of sodium acetate (3 M, 60 µL) and ethanol (96%, 1650 µL). After incubation of 1 h at -20°C, the solution was centrifuged (30 min, 15.000 g, 4°C). The pellet was washed with ethanol (70%) and resuspended in distilled water (30 µL). Prior to sequencing, the RNA was treated with DNase according to standard JGI protocols (http://jgi.doe.gov/collaborate-with-jgi/pmo-overview/protocols-sample-preparation-information/).

For preparation of stranded cDNA libraries for RNA-seq, mRNA was purified from 1 µg of total RNA using magnetic beads containing poly-T oligos. mRNA was fragmented using divalent cations and high temperature. The fragmented RNA was reverse transcribed using random hexamers and reverse transcriptase Superscript II (Invitrogen) followed by second strand synthesis. The fragmented cDNA was treated with end-pair, A-tailing, adapter ligation, and 10 cycles of PCR. qPCR was used to determine the concentration of the libraries. An overview of reads generated by paired-end sequencing is given in Text S1-Table S2.

b. Genome annotation

The genome assembly of *T. oleaginosus* was annotated using the JGI annotation pipeline (1), which combines several gene prediction and annotation methods. Before gene prediction, assembly scaffolds were masked using RepeatMasker (A.F.A. Smit, R. Hubley, P. Green, RepeatMasker Open-3.0, 1996-2010, http://www.repeatmasker.org) and RepBase library (2), with the most frequent (>150 times) repeats recognized by RepeatScout (3). The following combination of gene predictors was run on the masked assembly: *ab-initio* Fgenesh (4) and GeneMark (5), trained for specific genomes; homology-based Fgenesh+ (4) and Genewise (6), seeded by BLASTx alignments (7) against the NCBI-NR protein database; and transcriptome-based CombEST (Zhou et al., personal communication). In addition to protein-coding genes, tRNAs were predicted using tRNAscan-SE (8). All of the predicted proteins were functionally annotated using SignalP (9) for signal sequences, TMHMM (10) for transmembrane domains, InterProScan (11) for the integrated collection of functional and structured protein domains, and protein alignments to NCBI-NR, SwissProt (http://www.expasy.org/sprot/), KEGG (12) for metabolic pathways, and KOG (13) for eukaryotic clusters of orthologues. Interpro and SwissProt hits were also used to map the gene-ontology terms (14). For each genomic locus, the best representative gene model was

selected based on a combination of protein similarity and EST support. Completeness of the annotation was estimated using CEGMA 2.5 (15). *De novo* repeats were predicted with RepeatModeler (A.F.A. Smit, R. Hubley, www.repeatmasker.org/RepeatModeler.html) as described (16).

c. Quantitative analysis of gene expression

For quantitative analysis of gene expression, reads from each library were aligned to the reference genome using TopHat (17) with only unique mapping allowed. If a read mapped to more than one location, it was ignored. HTSeq (18) was used to generate the raw gene counts. Raw gene counts were used to evaluate the level of correlation between biological replicates using Pearson's correlation. All replicates showed high correlation coefficients (0.84-0.97) and were subsequently used for analysis of differential gene expression. DESeq2 (version 1.2.10) (19) was used to determine which genes were differentially expressed between pairs of conditions. The parameters used to call a gene differentially expressed between conditions were: adjusted p-value < 0.05. For an analysis of the 500 most strongly expressed genes in each condition (top500 analysis), RPKM (reads per kilobase per million mapped reads (20)) values were calculated for each gene and condition, and the genes with the 500 highest RPKM values for each condition were used for downstream analysis.

d. Phylogenetic analyses

Phylogenetic analyses were made with PAUP version 4.0b10 for Windows (D.L. Swofford, distributed by Sinauer Associates, copyright 2001 Smithsonian Institution) for DM and MP analyses, and with MrBayes (21). DM and MP analyses were performed as described using 1,000 bootstrap replicates, Bayesian analysis was performed with at least 250,000 generations (22). Consensus trees were graphically displayed with TREEVIEW (23). For generating a species tree of Tremellomycetes, clusters of orthologous genes among the fungal genomes used for the analyses were identified using OrthoMCL version 2.0.9 (24) using an inflation factor of 1.5. To reconstruct the phylogeny of these organisms, 200 orthologous groups of genes having exactly one gene in each organism were identified. The sequences of each species were concatenated, aligned using MAFFT version 7.123b (25) and well-aligned regions were extracted using Gblocks 0.91b (26). This resulted in 71463 amino acid positions. The parallelized version of RAxML version 8.1.16 (27) with the PROTGAMMAWAG model with 100 rapid bootstrap partitions was used to reconstruct a species tree. The tree was visualized using Dendroscope version 3.2.10 (28) and rooted on the outgroup *C. cinerea*.

e. Search for pheromone precursor genes

The assembled genome sequences of *T. oleaginosus* and *T. asahii* were screened for open reading frames (ORFs) of 15 to 55 amino acids encoding the conserved CAAX motif at the C-terminus (29) using custom-made Perl scripts. In addition, *T. oleaginosus* scaffolds 2, 68, and

70, which contain other mating type genes, were screened for ORFs up to 200 amino acids. Search results were checked for overlap with other annotated features and presence of conserved sequence residues besides the CAAX motif. To test whether putative pheromone genes might have been sequenced, but were not assembled, Illumina reads that did not map to the *T. oleaginosus* assembly using Bowtie2 (30) were searched by a k-mer based approach: All possible k-mers of 121 nt length were constructed from the unmapped reads, and k-mers with a coverage of at least 2 were screened for ORFs as described above using custom-made Perl scripts.

f. Analysis of mating type regions in different *T. oleaginosus* strains

DNA fragments containing mating type regions containing the *SXI1* and *STE3* genes were amplified by PCR from genomic DNA of strains ATCC20508 and ATCC20509 using the following primers: To_Sxi1_1 (CTCGCTTCGTTACTTCAAGGTCG) and To_Sxi1_2 (TCCGGAGATTCGCCGACGTTTGG) for *SXI1*, and To_Ste3_1 (GCAACCTCCCATTGACAGTCACC) and To_Ste3_4 (CGGTTTCCGTAACAACAACCAGC) for *STE3*. The ITS sequences were amplified using primers ITS1 (TCCGTAGGTGAACCTGCGG) and ITS4 (TCCTCCGCTTATTGATATGC) from the AFTOL project ((31), http://wasabi.lutzonilab.net/pub/primers/viewPrimers). PCR fragments were sequenced by Sanger sequencing.

2. Supplemental information about mating type analyses

a. Evolution of mating type regions in Tremellomycetes

Our analyses suggest that the most parsimonious assumption for the evolution of the mating type loci in the Tremellomycetes is an independent recruitment event for one set of developmental genes into the HD locus or P/R locus, respectively, and an independent loss of one HD gene in the *Trichosporon* and *Cryptococcus* lineages. This suggests that an evolutionary trend towards larger mating-type regions is present throughout, and that functionally similar results can be achieved by different genomic configurations. With respect to the loss of one HD gene, it was hypothesized that this might be an adaptation to a pathogenic lifestyle in the *Cryptococcus* species (32); however, *T. oleaginosus* is not pathogenic, and while *T. asahii* can occur as an opportunistic pathogen, this is not an obligate part of its life cycle. Thus, potential benefits of harboring only one HD transcription factor gene at the mating-type locus are unlikely to be related to pathogenicity in the *Trichosporon* species. A more general assumption is that fewer mating-specific genes reduce the potential for outbreeding or increase the possibility of inbreeding, thus allowing sexual development in situations where the population structure makes finding suitable mating partners difficult (32). However, whether such environmental restrictions exist for *T. oleaginosus* is not clear at present. Loss of function or actual loss of genes for potential key factors determining mating and regulating sexual development without apparent loss of a sexual cycle has been described for other fungi, e.g. *Coprinellus disseminatus* and several *Candida* species (33-36). In *Candida*, gene loss or loss of function has led to extensive rewiring of core meiotic processes, probably related to changes from a predominantly diploid to a haploid lifestyle in some species. Thus, while mating and meiosis are conserved, there is a high plasticity with respect to the molecular machinery that drives these core eukaryotic processes. This plasticity is reflected in the available genomic data for Tremellomycetes, with a consistent trend towards larger mating-type regions resulting in different, but presumably functionally similar genomic arrangements in different lineages.

b. Phylogenetic strata and alleles of mating type genes

In *C. neoformans* and other *Cryptococci*, genes at the mating-type locus can be grouped into several phylogenetically distinct strata, ranging from ancient mating-type genes like the HD transcriptions factors and the pheromone receptor gene *STE3*, which show a mating type-specific phylogenetic pattern, to genes that were more recently acquired into the mating-type locus and show a less distinct mating type-specific phylogeny or even a species-specific phylogeny (32, 37, 38). The *T. oleaginosus* Sxi1 protein homolog clusters with Sxi1 homologs in a phylogenetic analysis, suggesting a mating type-specific clustering, although confirmation of this finding would require the analysis of a predicted Sxi2 homolog, which is not present in any of the currently available *T. oleaginosus* strains (see below). The predicted homeodomain of Sxi1 shows the hallmarks of the HD1 class (Figure S1). Phylogenetic clustering suggesting mating type-specificity was also found for the Ste3 homolog, although it clusters with the *MAT***a**-specific Ste3 proteins from the *Cryptococci* (Figure S1), in contrast to

Sxi1, which clusters with *MATα* proteins (Figure S1). A multiple alignment of Ste3 homologs confirms this grouping, based on the absence of a conserved proline residue that is present in the *Cryptococci MATα* Ste3 proteins, but not in the MATa Ste3 proteins (Figure S2). Under the assumption that *T. oleaginosus* harbors two unlinked mating-type loci (tetrapolar mating system), any combination of HD transcription factor and pheromone receptor alleles is theoretically possible, in contrast to the bipolar *C. neoformans*, where only two combinations exist (39).

To test whether different alleles of the essential mating-type genes exist in the population, we analyzed the two previously described *T. oleaginosus* strains ATCC20508 and ATCC20509 (40). Genomic regions spanning *SXI1* and *STE3*, respectively, including non-coding upstream and downstream regions, were amplified by PCR and sequenced. In both strains, the *SXI1* and *STE3* genomic regions are identical to those from the sequenced strain IBC0246. Thus, the three *T. oleaginosus* strains carry the same mating-type configuration with respect to core mating-type genes, and at present it remains unknown how many other mating-type alleles and allele combinations exist in the population.

3. Text S1-Table 1. Media used for different growth conditions. Strains were propagated in complete medium (YPD obtained from Sigma-Aldrich, Steinheim, Germany) containing 2 % (w/v) glucose. For the analysis of transcriptomes and lipid production, strains were grown in YPD, NLM, PLM, NAG, or XYL media.

a) YPD medium (complete medium, Sigma-Aldrich, Steinheim, Germany)

Ingredients	**Concentration(g/L)**
Trypton	20
Yeast extract	10
Glucose	20

b) NLM: nitrogen limitation medium with xylose pH=6

Ingredients	**Concentration(g/L)**
Xylose	30
Yeast extract	0.75
$(NH_4)_2SO_4$	0.0012
$MgSO_4 \cdot 7H_2O$	1.5
KH_2PO_4	0.4
$CaCl_2 \cdot 2H_2O$	0.22
$ZnSO_4 \cdot 7H_2O$	0.55 µg L^{-1}
$MnCl_2 \cdot 4H_2O$	24.2 µg L^{-1}
$CuSO_4 \cdot 5H_2O$	25 µg L^{-1}

c) PLM: phosphate limitation medium with xylose pH=6

Ingredients	**Concentration(g/L)**
Xylose	30
Yeast extract	0.75
$(NH_4)_2SO_4$	4
$MgSO_4 \cdot 7H_2O$	1.5
KH_2PO_4	0.14
$CaCl_2 \cdot 2H_2O$	0.22
$ZnSO_4 \cdot 7H_2O$	0.55 µg L^{-1}
$MnCl_2 \cdot 4H_2O$	24.2 µg L^{-1}
$CuSO_4 \cdot 5H_2O$	25 µg L^{-1}

d) NAG: minimal medium with NAcGlc (2 %)

Ingredients	Concentration(g/L)
YNB (Sigma Aldrich)	6.7
NAcGlc	20

e) XYL: minimal medium with or xylose (2 %)

Ingredients	Concentration(g/L)
YNB (Sigma Aldrich)	6.7
Xylose	20

4. Text S1-Table 2. Overview of RNA-seq libraries. For each growth condition, three independent biological replicates were sequenced, with the exception of condition YPD, where only two independent replicates were sequenced.

condition	sampleName	no. of reads
NLM (nitrogen limitation)	T_ole_NLM_rep1	72,668,552
	T_ole_NLM_rep2	65,891,002
	T_ole_NLM_rep3	59,944,872
XYL (xylose as carbon source)	T_ole_XYL_rep1	30,405,872
	T_ole_XYL_rep2	33,461,342
	T_ole_XYL_rep3	77,656,936
PLM (phosphate limitation)	T_ole_PLM_rep1	30,790,406
	T_ole_PLM_rep2	29,659,850
	T_ole_PLM_rep3	34,608,594
NAG (N-acetyl glucosamin as carbon source)	T_ole_NAG_rep1	34,713,562
	T_ole_NAG_rep2	65,476,758
	T_ole_NAG_rep3	26,289,826
YPD (growth on complete medium)	T_ole_YPD_rep1	30,365,650
	T_ole_YPD_rep2	104,659,846

5. References for Text S1

1. **Grigoriev IV, Martines D, Salamov A.** 2006. Fungal genomic annotation. Appl Mycol Biotechnol **6:**123-142.
2. **Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J.** 2005. Repbase Update, a database of eukyrotic repetitive elements. Cytogenet Genome Res **110:**462-467.
3. **Price AL, Jones NC, Pevzner PA.** 2005. *De novo* identification of repeat families in large genomes. Bioinformatics **21:**i351-i358.
4. **Salamov AA, Solovyev VV.** 2000. Ab initio gene finding in Drosophila genomic DNA. Genome Res **10:**516-522.
5. **Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M.** 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Res **18:**1979-1990.
6. **Birney E, Clamp M, Durbin R.** 2004. GeneWise and Genomewise. Genome Res **14:**988-955.
7. **Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:**3389-3402.
8. **Lowe TM, Eddy SR.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucl Acids Res **25:**955-964.
9. **Petersen TN, Brunak S, von Heijne G, Nielsen H.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods **8:**785–786.
10. **Krogh A, Larsson B, von Heijne G, Sonnhammer ELL.** 2001. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J Mol Biol **305:**567-580.
11. **Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquenne L, Finn RD, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Laugraud A, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, Mistry J, Mitchell A, Mulder N, Natale D, Orengo C, Quinn AF, Selengut JD, Sigrist CJA, Thimma M, Thomas PD, Valentin F, Wilson D, Wu CH, Yeats C.** 2009. InterPro: the integrative protein signature database. Nucl Acids Res **37:**D211-215.
12. **Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y.** 2008. KEGG for linking genomes to life and the environment. Nucleic Acids Res **36:**D480-D484.
13. **Koonin EV, Fedorova ND, Jackson JD, Jacobs AR, Krylov DM, Makarova KS, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Rogozin IB, Smirnov S, Sorokin AV, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA.** 2004. A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. Genome Biol **5:**R7.
14. **Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G.** 2000. Gene Ontology: tool for the unification of biology. Nat Genet **25:**25-29.
15. **Parra G, Bradnam K, Ning Z, Keane T, Korf I.** 2009. Assessing the gene space in draft genomes. Nucl Acids Res **37:**289-297.
16. **Traeger S, Altegoer F, Freitag M, Gabaldon T, Kempken F, Kumar A, Marcet-Houben M, Pöggeler S, Stajich JE, Nowrousian M.** 2013. The genome and development-dependent transcriptomes of *Pyronema confluens*: a window into fungal evolution. PLoS Genet **9:**e1003820.

17. **Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg S.** 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol **14:**R36.
18. **Anders S, Pyl PT, Huber W.** 2014. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics doi:10.1093/bioinformatics/btu638.
19. **Love MI, Huber W, Anders S.** 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. bioRxiv **doi: http://dx.doi.org/10.1101/002832**.
20. **Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B.** 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods **5:**621-628.
21. **Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu LI, Suchard MA, Huelsenbeck JP.** 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol **61:**539-542.
22. **Hall BG.** 2004. Phylogenetic trees made easy, 2 ed. Sinauer Associates, Sunderland.
23. **Page R.** 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. Appl Biosci **12:**357-358.
24. **Li L, Stoeckert CJJ, Roos DS.** 2003. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. Genome Res **13:**2178-2189.
25. **Katoh K, Standley DM.** 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol **30:**772-780.
26. **Castresana J.** 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol **17:**540-552.
27. **Stamatakis A.** 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics **30:**1312-1313.
28. **Huson DH, Scornavacca C.** 2012. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. Syst Biol **61:**1061-1067.
29. **Kües U, James TY, Heitman J.** 2011. Mating type in basidiomycetes: unipolar, bipolar, and tetrapolar patterns of sexuality, p 97-160. *In* Pöggeler S, Wöstemeyer J (ed), The Mycota XIV Evolution of fungi and fungal-like organisms. Springer, Berlin, Heidelberg.
30. **Langmead B, Salzberg SL.** 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods **9:**357-359.
31. **James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J, Lumbsch HT, Rauhut A, Reeb V, Arnold AE, Amtoft A, Stajich JE, Hosaka K, Sung G-H, Johnson D, O/'Rourke B, Crockett M, Binder M, Curtis JM, Slot JC, Wang Z, Wilson AW, Schuszler A, Longcore JE, O/'Donnell K, Mozley-Standridge S, Porter D, Letcher PM, Powell MJ, Taylor JW, White MM, Griffith GW, Davies DR, Humber RA, Morton JB, Sugiyama J, Rossman AY, Rogers JD, Pfister DH, Hewitt D, Hansen K, Hambleton S, Shoemaker RA, Kohlmeyer J, Volkmann-Kohlmeyer B, Spotts RA, et al.** 2006. Reconstructing the early evolution of fungi using a six-gene phylogeny. Nature **443:**818-822.
32. **Heitman J, Sun S, James TY.** 2013. Evolution of fungal sexual reproduction. Mycologia **105:**1-27.
33. **James TY, Srivilai P, Kües U, Vilgalys R.** 2006. Evolution of the bipolar mating system of the mushroom *Coprinellus disseminatus* from its tetrapolar ancestors involved loss of mating-type-specific pheromone receptor function. Genetics **172:**1877-1891.
34. **Reedy JL, Floyd AM, Heitman J.** 2009. Mechanistic plasticity of sexual reproduction in the *Candida* pathogenic species complex. Curr Biol **19:**891-899.

35. **Butler G, Rasmussen MD, Lin MF, Santos MAS, Sakthikumar S, Munro CA, Rheinbay E, Grabherr M, Forche A, Reedy JL, Agrafioti I, Arnaud MB, Bates S, Brown AJP, Brunke S, Costanzo MC, Fitzpatrick DA, de Groot PWJ, Harris D, Hoyer LL, Hube B, Klis FM, Kodira C, Lennard N, Logue ME, Martin R, Neiman AM, Nikolaou E, Quail MA, Quinn J, Santos MC, Schmitzberger FF, Sherlock G, Shah P, Silverstein KAT, Skrzypek MS, Soll D, Staggs R, Stansfield I, Stumpf MPH, Sudbery PE, Srikantha T, Zeng Q, Berman J, Berriman M, Heitman J, Gow NAR, Lorenz MC, Birren BW, Kellis M, et al.** 2009. Evolution of pathogenicity and sexual reproduction in eight *Candida* genomes. Nature **459:**657-662.
36. **Sherwood RK, Scaduto CM, Torres SE, Bennett RJ.** 2014. Convergent evolution of a fused sexual cycle promotes the haploid lifestyle. Nature **506:**387-390.
37. **Findley K, Sun S, Fraser JA, Hsueh YP, Averette AF, Li W, Dietrich FS, Heitman J.** 2012. Discovery of a modified tetrapolar sexual cycle in *Cryptococcus amylolentus* and the evolution of *MAT* in the Cryptococcus species complex. PLoS Genet **8:**e1002528.
38. **Metin B, Findley K, Heitman J.** 2010. The mating type locus (*MAT*) and sexual reproduction of *Cryptococcus heveanensis*: insights into the evolution of sex and sex-determining chromosomal regions in fungi. PLoS Genet **6:**e1000961.
39. **Hsueh YP, Heitman J.** 2008. Orchestration of sexual reproduction and virulence by the fungal mating-type locus. Curr Opin Microbiol **11:**517-524.
40. **Gujjari P, Suh SO, Coumes K, Zhou JJ.** 2011. Characterization of oleaginous yeasts revealed two novel species: *Trichosporon cacaoliposimilis* sp. nov. and *Trichosporon oleaginosus* sp. nov. Mycologia **103:**1110-1118.